

Matrix Decompositions using sub-Gaussian Random Matrices

Yariv Aizenbud¹ Amir Averbuch²

¹School of Applied Mathematics, Tel Aviv University, Israel

²School of Computer Science, Tel Aviv University, Israel

February 11, 2016

Abstract

In recent years, several algorithms, which approximate matrix decomposition, have been developed. These algorithms are based on metric conservation features for linear spaces of random projection types. We show that an i.i.d sub-Gaussian matrix with large probability to have zero entries is metric conserving. We also present a new algorithm, which achieves with high probability, a rank r decomposition approximation for an $m \times n$ matrix that has an asymptotic complexity like state-of-the-art algorithms. We derive an error bound that does not depend on the first r singular values. Although the proven error bound is not as tight as the state-of-the-art bound, experiments show that the proposed algorithm is faster in practice, while getting the same error rates as the state-of-the-art algorithms get.

Keywords. SVD decomposition, LU decomposition, Low rank approximation, random matrices, sparse matrices, sub-Gaussian matrices, Johnson-Lindenstrauss Lemma, oblivious subspace embedding.

1 Introduction

Dimensionality reduction by randomized linear maps preserves metric features. The Johnson-Lindenstrauss Lemma (JL) [11] shows that there is a random distribution of linear dimensionality reduction operators that preserves, with bounded error and high probability, the norm of a set of vectors. For example, Gaussian random matrices satisfy this property.

JL Lemma was extended in the following way. While the classical formulation dealt with norm conservation of sets of vectors, the JL-based extension deals with a subspace of a vector space. This extension is considered for example in [22], where it shows that Fourier based random matrices of size $n \times \mathcal{O}(r \log r)$ conserves the norm of all the vectors from a vector space of dimension r . Similar results for sparse matrices distribution are given in [3, 5, 12, 15].

In recent years, several algorithms that approximate matrix decomposition, which are based on norm conservation, have been developed. The idea is roughly as follows: A randomly drawn matrix Ω , which projects the original matrix into a lower dimension, is used. The decomposition is calculated in the low dimensional space. Then, this decomposition is mapped into the matrix original size. It is shown in [14, 20] how to use random Gaussian matrices in order to find, with high probability, an approximated interpolative decomposition, singular value decomposition (SVD) and LU decomposition. FFT-based random matrices, which approximate matrix decompositions, are described in [24]. The special structure of the FFT-based distribution provides a fast matrix multiplication that yields a faster algorithm than the algorithms in [14]. A comprehensive review of these ideas (and many more) is given in [9]. The algorithm in [3] uses a sparse random matrix distribution that makes the matrix multiplication step in the algorithm even faster than what the FFT-based matrices provide.

In this paper, we show that the class of matrices with i.i.d sub-Gaussian entries satisfy the image conservation property even when the probability for a zero entry grows with the size of the matrix. Additionally, we construct fast SVD and LU decomposition algorithms with bounded error and asymptotic complexity equal to the asymptotic complexity of the state-of-the-art algorithm. Although the asymptotic complexity is the same, the practical running time of the presented algorithms is lower than the existing algorithms. Since the random projections are matrices with i.i.d entries,

it is not required to set the dimension k of the projection in advance. It is possible, although not elaborated in this paper, to increase k iteratively, until the resulting approximation is in the required accuracy. Stronger bounds for the case of sparse-Bernoulli random matrices are shown in [4]¹

We denote by $M_{n \times m}$ the set of n by m matrices. We call a rectangular random matrix distribution \mathcal{M} an *metric conserving* distribution if for any $A \in M_{n \times m}$ a randomly chosen $\Omega \in M_{m \times k}$ from \mathcal{M} , the image of $A\Omega$ is similar to the image of A . Three main parameters related to this property are the dimension k of Ω (the smaller the better), the “distance” between the images of $A\Omega$ and A and the probability for which the image conservation is valid. It is obvious that these parameters are connected. Distributions, which conserve the norm allowing an error $(1 + \varepsilon)$ of the theoretical bound, are called *oblivious subspace embedding (OSE)* ([15]).

The theoretical bound for a rank r approximation of a matrix A in L_2 norm is $\sigma_{r+1}(A)$ and in Frobenius norm it is Δ_{r+1} , where $\sigma_r(A)$ is the r th largest singular value of A and $\Delta_r(A) \triangleq (\sum_{l=r}^n \sigma_l^2(A))^{1/2}$. Three important results related to the above parameters, which deal with metric conserving distributions in the context of randomized decomposition algorithms, are: 1. Achieving an accuracy of $\mathcal{O}_\sigma(\sigma_{r+1}(A))$ for a rank r measured in L_2 norm with high probability, is described in [9, 14]. To achieve this accuracy with high probability, the required Ω can be an i.i.d Gaussian matrix of size $\mathcal{O}(r)$. 2. Achieving an accuracy of $\mathcal{O}_\sigma(\sigma_{r+1}(A))$ for a rank r measured in L_2 norm with high probability, is described in [9, 24]. To achieve this accuracy with high probability, Ω can be an FFT-based matrix of size $\mathcal{O}(r \log r)$. 3. The result in [15] achieves accuracy of $(1 + \varepsilon)\Delta_{r+1}(A)$ with high probability measured in Frobenius norm. While Ω is drawn from a sparse distribution, its size is assumed to be not less than $\mathcal{O}(r^2/\varepsilon^2)$. In fact, for sparse matrices distribution, a lower bound for the size of Ω is provided in [16].

We show in Section 3 that for the class of matrices with i.i.d sub-Gaussian entries, the size of Ω , which is needed to achieve an accuracy $\mathcal{O}_\sigma(\sigma_{r+1}(A))$ measured in L_2 norm. We also show its dependency on the probability to have a zero entry. By choosing a sparse matrix distribution to be sub-Gaussian, we were able to perform a fast matrix multiplication while having a small size Ω . It is shown in [6] that this class of sub-Gaussian matrices of size $\mathcal{O}(r/\varepsilon^2)$ with constant probability distribution is an OSE. In this paper, we provide a bound for the case where the distribution depends on the size of the matrix.

The state-of-the-art result for rank r approximation algorithm appears in [3]. It describes how to use a sparse embedding matrix to construct an algorithm that finds for any matrix $A \in M_{m \times n}$ and any rank r , with high probability, an SVD approximation of rank r . Namely, orthogonal U, V^* and a diagonal matrix Σ are formed such that $\|A - U\Sigma V^*\|_F \leq (1 + \varepsilon)\Delta_{r+1}(A)$. Although the algorithm in [9] uses a smaller Ω than [3], the algorithm in [3] is asymptotically faster than the algorithm in [9] because of the sparse nature of the projection.

We describe in Section 4.1 an algorithm that for each $A \in M_{m \times n}$ outputs with high probability a low rank SVD approximation that is built from U, Σ and V . The algorithm works with any metric conserving or OSE random distribution. The size k of the random embedding in the algorithm depends on the probability p for having a zero entry. The complexity of the algorithm when using i.i.d sub-Gaussian random matrix projections is $\mathcal{O}(\text{nnz}(A)pk + (m + n)k^2)$ where $\text{nnz}(A)$ denotes the number of non-zeros in A and $k = \mathcal{O}(\frac{1}{p^3} \ln r)$. For sparse embedding matrix distribution as in [15], the complexity of the algorithm in Section 4.1 is the same as in [15]. This algorithm guarantees with high probability that $\|A - U\Sigma V^*\|_2 \leq \mathcal{O}_\sigma(\sigma_{r+1}(A))$. Although the guaranteed error bound is less tight than the one in [3], we show in Section 5 that in practice our algorithm reaches the same error in less time.

The randomized LU decomposition algorithm in [1] is based on the ideas from [3]. We show in Section 4.2 that it is also valid when random matrices from a sub-Gaussian distribution are chosen with the complexity and error bound equal to those from the SVD decomposition.

The paper has the following structure: In Section 2, we present the necessary mathematical preliminaries. In Section 3, we show that i.i.d sub-Gaussian random matrices are metric conserving and in Section 4 we describe the SVD algorithm and show that the LU algorithm in [1] is valid with i.i.d sub-Gaussian random matrices. In section 5, we present the numerical results of the described SVD algorithm.

¹ The results on sub-Gaussian random matrices in this paper were derived couple of months before the paper of [4] was brought to our attention.

2 Preliminaries

2.1 The ε -Net

ε -net is defined in Definition 2.1. Its size is bounded by Lemma 2.1 that is proved in [19]. Throughout the paper, S^{n-1} denotes the $(n-1)$ -sphere in \mathbb{R}^n .

Definition 2.1. Let (T, d) be a metric space and let $K \subset T$. A set $\mathcal{N} \subset T$ is called ε -net of K if for all $x \in K$ there exists $y \in \mathcal{N}$ such that $d(x, y) < \varepsilon$.

Lemma 2.1 (Proposition 2.1 in [19]). For any $\varepsilon < 1$, there exists an ε -net \mathcal{N} of S^{n-1} such that

$$|\mathcal{N}| \leq 2n \left(1 + \frac{2}{\varepsilon}\right)^{n-1}.$$

Remark. It follows that for sufficiently large n , the size of $1/2$ -net of S^{n-1} has at most

$$2n \left(1 + \frac{2}{1/2}\right)^{n-1} = 2n \cdot 5^{n-1} \leq 6^n$$

points.

2.2 Compressible and Incompressible Vectors

Definition 2.2. A vector $v \in \mathbb{R}^n$ is called (η, ε) -incompressible if $\sum_{j: |v_j| \leq \varepsilon} |v_j|^2 \geq \eta^2$ and compressible otherwise.

Lemma 2.2. Let $U \subset \mathbb{R}^n$ a subspace of dimension r . Let \mathcal{N} be an ε_{net} -net of the set of (η, ε_c) -compressible vectors in U . Then,

$$|\mathcal{N}| \leq r^{\frac{1}{\varepsilon_c^2}} \eta^{r - \frac{1}{\varepsilon_c^2}} \left(\frac{c_{net}}{\varepsilon_{net}} \right)^r$$

for an absolute constant c_{net} .

Proof. The (η, ε_c) -compressible vectors are in an η distance from a sparse vector with no more than $\frac{1}{\varepsilon_c^2}$ non-zero coordinates. For small enough η , the volume of η -balls around $\frac{1}{\varepsilon_c^2}$ -sparse vectors is $r^{\frac{1}{\varepsilon_c^2}} \eta^{r - \frac{1}{\varepsilon_c^2}}$. The same arguments from the proof of Lemma 2.1 show that the number of points in an ε_{net} -net of this volume is not more than $r^{\frac{1}{\varepsilon_c^2}} \eta^{r - \frac{1}{\varepsilon_c^2}} \left(\frac{c_{net}}{\varepsilon_{net}} \right)^r$. \square

2.3 Sub-Gaussian Random Variables

In this section, we introduce the sub-Gaussian random variables with some of their properties. Sub-Gaussian variables are an important class of random variables that have strong tail decay properties. This class contains, for example, all the bounded random variables and the normal variables.

Definition 2.3. A random variable X is called sub-Gaussian if there exists constants v and C such that for any $t > 0$, $\mathbb{P}(|X| > t) \leq Ce^{-vt^2}$ and X has a non-zero variance. A random variable X is called centered if $\mathbb{E}X = 0$.

Remark. For convenience, we use the term *sub-Gaussian matrix* for a matrix with i.i.d sub-Gaussian entries.

Many non-asymptotic results on a sub-Gaussian matrix distribution have recently appeared. A survey of this topic appears in [17, 23].

The following facts, proved in [13, 17–19, 23], are used in the paper:

1. Linear combination of centered sub-Gaussian variables is also sub-Gaussian. This is stated in Theorem 2.3. The inequality in this theorem is similar to Hoeffding inequality [10].
2. The bound for the first singular value of a sub-Gaussian random matrix is given in Theorem 2.4.

3. The probability bound for the sum of centered sub-Gaussian variables to be small is given in Theorem 2.6.

Formally,

Theorem 2.3. *Let X_1, \dots, X_n be independent centered sub-Gaussian random variables. Then, for any $a_1, \dots, a_n \in \mathbb{R}$*

$$\mathbb{P} \left(\left| \sum_{j=1}^n a_j X_j \right| > t \right) \leq 2 \exp \left(- \frac{ct^2}{\sum_{j=1}^n a_j^2} \right).$$

Theorem 2.4. *Let Ω be a $k \times n$, $n \geq k$, random matrix whose entries are i.i.d centered sub-Gaussian random variable. Then, $\mathbb{P}(\sigma_1(\Omega) > t\sqrt{n}) \leq e^{-ct^2n}$ holds for $t \geq C_0$.*

Since we are interested in sparse matrices, the following definition is useful.

Definition 2.4. *A sub-Gaussian random variables X is represented by a combination of a centered sub-Gaussian random variable $\frac{1}{\sqrt{p}}Z$ with $\mathbb{P}(Z=0)=0$, $\mathbb{E}(Z^2)=1$ with probability p and 0 otherwise.*

Note that $\mathbb{E}(X)=0$, $\mathbb{E}(X^2)=1$, $\mathbb{E}(X^3)=\frac{\mathbb{E}(Z^3)}{\sqrt{p}}$ and $\mathbb{E}(X^4)=\frac{\mathbb{E}(Z^4)}{p}$.

Lemma 2.5. *Let X_1, \dots, X_n be independent centered sub-Gaussian random variables defined as a combination of a centered sub-Gaussian $\frac{1}{\sqrt{p}}Z$ with $\mathbb{P}(Z=0)=0$ and $\mathbb{E}(Z^2)=1$ with probability p and 0 otherwise. Then, for any $(a_1 \dots a_n) \in S^{n-1}$ the third and forth moment (skewness and kortosis) of $(\sum_{i=1}^n a_i X_i)$ are bounded by*

$$\mathbb{E} \left(\left(\sum_{i=1}^n a_i X_i \right)^3 \right) \leq \frac{\mathbb{E}(Z^3)}{\sqrt{p}}$$

and

$$\mathbb{E} \left(\left(\sum_{i=1}^n a_i X_i \right)^4 \right) \leq \frac{\mathbb{E}(Z^4) + 1}{p} \triangleq \frac{z_4}{p}.$$

Proof.

$$\begin{aligned} \mathbb{E} \left(\left(\sum_{i=1}^n a_i X_i \right)^3 \right) &= \sum_{i=1}^n a_i^3 \mathbb{E}(X_i^3) + \sum_{i,j=1, i \neq j}^n a_i^2 a_j \mathbb{E}(X_i^2) \mathbb{E}(X_j) \leq \mathbb{E}(X^3) = \frac{\mathbb{E}(Z^3)}{\sqrt{p}}. \\ \mathbb{E} \left(\left(\sum_{i=1}^n a_i X_i \right)^4 \right) &= \sum_{i=1}^n a_i^4 \mathbb{E}(X_i^4) + \sum_{i,j=1, i \neq j}^n a_i^2 a_j^2 \leq \mathbb{E}(X^4) + 1 = \frac{\mathbb{E}(Z^4) + p}{p}. \end{aligned}$$

Since $p \leq 1$, the proof is completed. \square

Lemma 2.6. *Let X_1, \dots, X_n be an i.i.d centered sub-Gaussian random variable as in Definition 2.4. For every coefficients vector (in particular for a compressible vector) $a = (a_1, \dots, a_n) \in S^{n-1}$, the random sum $S = \sum_{i=1}^n a_i X_i$ satisfies $\mathbb{P}(|S| < \lambda) \leq 1 - p \frac{(1-\lambda^2)^2}{z_4}$.*

Proof. Let $0 < \lambda < (\mathbb{E}S^2)^{1/2} = 1$. By the Cauchy-Schwarz inequality,

$$\mathbb{E}S^2 = \mathbb{E}S^2 \mathbf{1}_{[-\lambda, \lambda]}(S) + \mathbb{E}S^2 \mathbf{1}_{\mathbb{R} \setminus [-\lambda, \lambda]}(S) \leq \lambda^2 + (\mathbb{E}S^4)^{1/2} \mathbb{P}(|S| > \lambda)^{1/2}.$$

This leads to the Paley-Zygmund inequality:

$$\mathbb{P}(|S| > \lambda) \geq \frac{(\mathbb{E}S^2 - \lambda^2)^2}{\mathbb{E}S^4} = \frac{(1 - \lambda^2)^2}{\mathbb{E}S^4}.$$

By Theorem 2.3, the random variable S is sub-Gaussian. By Lemma 2.5, $\mathbb{E}S^4 \leq \frac{z_4}{p}$ where $z_4 = \mathbb{E}Z^4 + 1$. To complete the proof

$$\mathbb{P}(|S| < \lambda) \leq 1 - \frac{(1 - \lambda^2)^2}{\mathbb{E}S^4} = 1 - p \frac{(1 - \lambda^2)^2}{z_4}.$$

In particular, for $\lambda = 1/2$ we have $\mathbb{P}(|S| < 1/2) \leq 1 - z_4' p$ for $z_4' = \frac{9}{16z_4}$

\square

Lemma 2.7. For any $0 < \alpha < 1$, there is c_s such that for any k , $\binom{k}{\alpha k} < c_s^{\alpha k \ln(\frac{1}{\alpha} - 1)} < c_s^{\alpha k \ln \frac{1}{\alpha}}$.

Proof. We use the Stirling formula to estimate $\ln \binom{k}{\alpha k}$.

$$\begin{aligned}
\ln \binom{k}{\alpha k} &= k \ln k - k - (\alpha k \ln(\alpha k) - \alpha k) - ((k - \alpha k) \ln(k - \alpha k) - (k - \alpha k)) + \mathcal{O}(\ln k) \\
&= k \ln k - \alpha k \ln(\alpha k) - k \ln(k - \alpha k) + \alpha k \ln(k - \alpha k) + \mathcal{O}(\ln k) \\
&= k \ln k - \alpha k \ln \alpha k - k \ln k(1 - \alpha) + \alpha k \ln(\alpha k(\frac{1}{\alpha} - 1)) + \mathcal{O}(\ln k) \\
&= \alpha k \ln(\frac{1}{\alpha} - 1) - k \ln(1 - \alpha) + \mathcal{O}(\ln k) \\
&\sim \alpha k \ln(\frac{1}{\alpha} - 1) - k(-\alpha - \frac{\alpha^2}{2} - \dots) \sim \alpha k \ln(\frac{1}{\alpha} - 1).
\end{aligned}$$

□

Lemma 2.8 follows from Berry-Essen's theorem [2, 7] in a similar fashion to the derivations in [21].

Lemma 2.8. For $S = \sum_{i=1}^n a_i X_i$ where X_i are i.i.d random variables with $E(X) = 0, E(X^2) = 1$ and $\sum_{i=1}^n a_i^2 = 1$, then for all r

$$\sup_t \mathbb{P} \left(\left| \sum_{i=1}^n a_i X_i - t \right| < r \right) \leq \frac{r}{\sqrt{\sum_{i=1}^n a_i^2}} + 2C_{BE} \frac{\mathbb{E}(X^3) \sum_{i=1}^n |a_i|^3}{\left(\sum_{i=1}^n a_i^2 \right)^{3/2}} \quad (2.1)$$

holds.

Proof. Let N be a standard normal variable. From Berry-Essen's theorem follows that for all r

$$\left| \mathbb{P} \left(\frac{\sum_{i=1}^n a_i X_i}{\sqrt{\sum_{i=1}^n a_i^2}} < \frac{r}{\sqrt{\sum_{i=1}^n a_i^2}} \right) - \mathbb{P} \left(N < \frac{r}{\sqrt{\sum_{i=1}^n a_i^2}} \right) \right| \leq C_{BE} \frac{\mathbb{E}(X^3) \sum_{i=1}^n |a_i|^3}{(\sum_{i=1}^n a_i^2)^{2/3}}.$$

Thus, for any t ,

$$\begin{aligned}
&\left| \mathbb{P} \left(\frac{|\sum_{i=1}^n a_i X_i - t|}{\sqrt{\sum_{i=1}^n a_i^2}} < \frac{r}{\sqrt{\sum_{i=1}^n a_i^2}} \right) - \mathbb{P} \left(|N - t| < \frac{r}{\sqrt{\sum_{i=1}^n a_i^2}} \right) \right| \\
&= \left| \mathbb{P} \left(\frac{\sum_{i=1}^n a_i X_i}{\sqrt{\sum_{i=1}^n a_i^2}} < \frac{t+r}{\sqrt{\sum_{i=1}^n a_i^2}} \right) - \mathbb{P} \left(N < \frac{t+r}{\sqrt{\sum_{i=1}^n a_i^2}} \right) - \right. \\
&\quad \left. \mathbb{P} \left(\frac{\sum_{i=1}^n a_i X_i}{\sqrt{\sum_{i=1}^n a_i^2}} < \frac{t-r}{\sqrt{\sum_{i=1}^n a_i^2}} \right) - P \left(N < \frac{t-r}{\sqrt{\sum_{i=1}^n a_i^2}} \right) \right| \quad (2.2) \\
&\leq 2C_{BE} \frac{\mathbb{E}(X^3) \sum_{i=1}^n |a_i|^3}{(\sum_{i=1}^n a_i^2)^{3/2}}.
\end{aligned}$$

By rewriting 2.2, we have

$$\mathbb{P} \left(\left| \sum_{i=1}^n a_i X_i - t \right| < r \right) \leq \mathbb{P} \left(|N - t| < \frac{r}{\sqrt{\sum_{i=1}^n a_i^2}} \right) + 2C_{BE} \frac{\mathbb{E}(X^3) \sum_{i=1}^n |a_i|^3}{(\sum_{i=1}^n a_i^2)^{3/2}}. \quad (2.3)$$

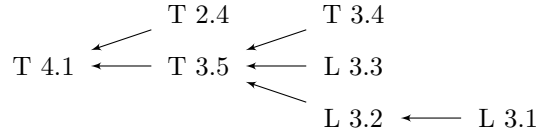
For any t , $\mathbb{P} \left(|N - t| < \frac{r}{\sqrt{\sum_{i=1}^n a_i^2}} \right) \leq \mathbb{P} \left(|N| < \frac{r}{\sqrt{\sum_{i=1}^n a_i^2}} \right) < \frac{r}{\sqrt{\sum_{i=1}^n a_i^2}}$. □

3 Metric conservation of sub-Gaussian random matrices

The main goal of this section is to show that for any matrix A and for a sub-Gaussian matrix Ω , the image of $A\Omega$ is “close” to the image of A with high probability, or, in other words, Ω preserves the geometry. Namely, if Q is an orthogonal basis for $A\Omega$, then $\|A - QQ^*A\|_2$ is small.

In order to show that the application of a random sub-Gaussian matrix preserves the geometry of A , we have to bound its behavior in any subspace of a given dimension r . We show in Theorem 3.4 that the norm of a random sub-Gaussian matrix in a subspace of dimension r is bounded from above with high probability. In Lemmas 3.2 and 3.3 it is shown that Ω conserves compressible and incompressible vectors, respectively, from a subspace of dimension r . In Theorem 3.5, these results are joined to show that the minimal singular value is bounded from below with high probability. The flow of the proof is based on ideas from the proof of bounds on singular values of Bernoulli random matrix in [21] and ideas from [17]. In Theorem 4.1, these results and the fact that the norm of a random matrix is also bounded (Theorem 2.4) are used to show that a sub-Gaussian matrix preserves the geometry.

These are the dependencies among the different theorems in this section:



Lemma 3.1. *Let X_1, \dots, X_n be i.i.d centered sub-Gaussian random variables as in Definition 2.4. Denote $\varepsilon_c = \varepsilon_0 \eta \sqrt{p}$. For any (η, ε_c) -incompressible $a = (a_1, \dots, a_n) \in S^{n-1}$, the random sum $\sum_{i=1}^n a_i X_i$ satisfies*

$$P(|\sum_{i=1}^n a_i X_i| \leq 2C\varepsilon_0\eta) \leq 2C_1(C)\varepsilon_0\mathbb{E}(Z^3)$$

where C is a constant that will be chosen later, and $C_1(C)$ depends only on C .

Proof. We recall that a is incompressible if $\sum_{j:|a_j| \leq \varepsilon_c} |a_j|^2 \geq \eta^2$. By using Lemma 2.8 we get:

$$\sup_t \mathbb{P}(|\sum_{i=1}^n a_i X_i - t| \leq r) \leq \frac{r}{\sqrt{\sum_{i=1}^n a_i^2}} + 2C_{BE} \frac{\mathbb{E}(X^3) \sum_{i=1}^n |a_i|^3}{\left(\sum_{i=1}^n a_i^2\right)^{3/2}}.$$

Note that we can condition out variables,

$$\sup_t \mathbb{P}(|\sum_{i=1}^n a_i X_i - t| \leq r) \leq \sup_t \mathbb{P}(|\sum_{i=1}^n a_i X_i - t| \leq r | X_1 = x_1).$$

If we condition out all the X_i for which $a_i > \varepsilon_c$, we get

$$\sup_t \mathbb{P}(|\sum_{i=1}^n a_i X_i - t| \leq r) \leq \frac{r}{\sqrt{\sum_{j:|a_j| \leq \varepsilon_c} a_j^2}} + 2C_{BE} \frac{\mathbb{E}(X^3) \sum_{j:|a_j| \leq \varepsilon_c} |a_j|^3}{\left(\sum_{j:|a_j| \leq \varepsilon_c} a_j^2\right)^{3/2}} \leq \frac{r}{\eta} + 2C_{BE} \frac{\mathbb{E}(X^3)\varepsilon_c}{\eta}.$$

By substituting $r = 2C\varepsilon_0\eta$ we have

$$\sup_t \mathbb{P}(|\sum_{i=1}^n a_i X_i - t| \leq 2C\varepsilon_0\eta) \leq \frac{2C\varepsilon_0\eta}{\eta} + 2C_{BE} \frac{\mathbb{E}(X^3)\varepsilon_c}{\eta}.$$

By using Lemma 2.5 and by substituting $\varepsilon_c = \varepsilon_0 \eta \sqrt{p}$ the proof is completed. \square

Lemma 3.2 (Ω conserves incompressible vectors in a subspace). *Let Ω be a $k \times n$ ($n \geq k$) random matrix whose entries are i.i.d centered sub-Gaussian random variable as in Definition 2.4. Denote $\varepsilon_c = \varepsilon_0 \eta \sqrt{p}$. Then, for any (ε_c, η) -incompressible $x \in S^{n-1}$,*

$$\mathbb{P}(\|\Omega x\|_2 < 2C\alpha\varepsilon_0\eta\sqrt{k}) \leq (C_1(C)\varepsilon_0\mathbb{E}(Z^3))^{k/4}$$

for a constant α .

Proof. The coordinates of the vector Ωx are independent linear combinations of i.i.d. sub Gaussian random variables with incompressible coefficients $(x_1 \dots x_n) \in S^{n-1}$. Hence, by Lemma 3.1, $\mathbb{P}(|(\Omega x)_j| < 2C\varepsilon_0\eta) \leq C_1(C)\varepsilon_0\mathbb{E}(Z^3) = \mu$ for all $j = 1, \dots, N$.

Assume that $\|\Omega x\|_2 < 2C\varepsilon_0\eta\alpha\sqrt{k}$. Then, $|(\Omega x)_j| < 2C\varepsilon_0\eta$ for at least $\lfloor (1 - \alpha^2)k \rfloor$ coordinates. Thus ,

$$\begin{aligned} \mathbb{P}(\|\Omega x\|_2 < 2C\varepsilon_0\eta\alpha\sqrt{k}) &\leq \mathbb{P}(\text{at least } \lfloor (1 - \alpha^2)k \rfloor \text{ coordinates satisfy } |(\Omega x)_j| < 2C\varepsilon_0\eta) \\ &= \sum_{l < \lfloor (1 - \alpha^2)k \rfloor}^k \binom{k}{l} \mathbb{P}(|(\Omega x)_1| < 2C\varepsilon_0\eta)^l (1 - \mathbb{P}(|(\Omega x)_1| < 2C\varepsilon_0\eta))^{k-l} \\ &\leq \sum_{l < \lfloor (1 - \alpha^2)k \rfloor}^k \binom{k}{l} \mu^l \\ &\leq \mu^{\lfloor (1 - \alpha^2)k \rfloor} \sum_{l < \lfloor (1 - \alpha^2)k \rfloor}^k \binom{k}{l}. \end{aligned}$$

If α is sufficiently small, then $\lfloor (1 - \alpha^2)k \rfloor > k/2$ and

$$\mathbb{P}(\|\Omega x\|_2 < 2C\varepsilon_0\eta\alpha\sqrt{k}) \leq \mu^{k/2} \alpha^2 k \binom{k}{\lfloor (1 - \alpha^2)k \rfloor}.$$

For α sufficiently small, $\alpha^2 k \binom{k}{\lfloor (1 - \alpha^2)k \rfloor} \leq \mu^{-k/4}$. Thus, $\mathbb{P}(\|\Omega x\|_2 < 2C\varepsilon_0\eta\alpha\sqrt{k}) \leq \mu^{-k/4} \cdot \mu^{k/2} \leq \mu^{k/4}$. \square

Lemma 3.3 (Ω conserves any vector in a subspace). *Let Ω be a $k \times n$ ($n \geq k$) random matrix whose entries are i.i.d centered sub-Gaussian random variable with variance 1 as in Definition 2.4. Let C be a constant that will be chosen later, and let η be small enough, such that $\eta^2 \ln \frac{1}{\eta} < c_4 p$. Then, for every $x \in S^{n-1}$, $\mathbb{P}(\|\Omega x\|_2 < 2C\eta\sqrt{k}) \leq (1 - z'_4 p)^{k/4}$ holds for a constant c_4 .*

Proof. The coordinates of the vector Ωx are independent linear combinations of i.i.d. sub Gaussian random variables with coefficients $(x_1 \dots x_n) \in S^{n-1}$. Hence, for $\lambda = 1/2$, by Lemma 2.6 $\mathbb{P}(|(\Omega x)_j| < 1/2) \leq 1 - z'_4 p$, $j = 1, \dots, N$.

Assume that $\|\Omega x\|_2 < 2C\eta\sqrt{k}$. Then, $|(\Omega x)_j| < 1/2$ for at least $\lfloor (1 - 4 \cdot 2^2 C^2 \eta^2)k \rfloor$ coordinates. Thus

$$\begin{aligned} \mathbb{P}(\|\Omega x\|_2 < 2C\eta\sqrt{k}) &\leq \mathbb{P}(\# \lfloor (1 - 16C^2\eta^2)k \rfloor \text{ coordinates satisfy } |(\Omega x)_j| < 1/2) \\ &= \sum_{l < \lfloor (1 - 16C^2\eta^2)k \rfloor}^k \binom{k}{l} \mathbb{P}(|(\Omega x)_1| < 1/2)^l (1 - \mathbb{P}(|(\Omega x)_1| < 1/2))^{k-l} \\ &\leq \sum_{l < \lfloor (1 - 16C^2\eta^2)k \rfloor}^k \binom{k}{l} (1 - z'_4 p)^l \\ &\leq (1 - z'_4 p)^{\lfloor (1 - 16C^2\eta^2)k \rfloor} \sum_{l < \lfloor (1 - 16C^2\eta^2)k \rfloor}^k \binom{k}{l}. \end{aligned}$$

If η is sufficiently small, then $\lfloor (1 - 16C^2\eta^2)k \rfloor > k/2$ and

$$\mathbb{P}(\|\Omega x\|_2 < 2C\eta\sqrt{k}) \leq (1 - z'_4 p)^{k/2} 16C^2 k \binom{k}{\lfloor (1 - 16C^2\eta^2)k \rfloor}.$$

From Lemma 2.7 follows that for η sufficiently small,

$$16C^2 k \binom{k}{\lfloor (1 - 16C^2\eta^2)k \rfloor} \leq 16C^2 k c_s^{16C^2\eta^2 k \ln(\frac{1}{16C^2\eta^2} - 1)} < c_1^{\eta^2 k \ln(\frac{1}{16C^2\eta^2} - 1)} < c_2^{k\eta^2 \ln \frac{1}{\eta^2}}.$$

Additionally, $(1 - z'_4 p)^{-k/4} > c_3^{c_2 p k}$. Thus, for η such that $\eta^2 \ln \frac{1}{\eta} < c_4 p$, $16C^2 k \binom{k}{\lfloor (1 - 16C^2\eta^2)k \rfloor} \leq (1 - z'_4 p)^{-k/4}$ holds. Thus, $\mathbb{P}(\|\Omega x\|_2 < 2C\eta\sqrt{k}) \leq (1 - z'_4 p)^{-k/4} \cdot (1 - z'_4 p)^{k/2} \leq (1 - z'_4 p)^{k/4}$. \square

The proof of Theorem 3.4 is similar to the proof of Theorem 2.4.

Theorem 3.4 (Maximum value in a subspace). *Let $U \subset \mathbb{R}^n$ be a linear subspace of dimension r . Let Ω be a $k \times n$ random matrix where $n \geq k > r$ and $k = \mathcal{O}(r)$ is sufficiently large. Assume the entries of Ω are i.i.d centered sub-Gaussian random variables. Then, for $t \geq C_0$ we have*

$$\mathbb{P} \left(\max_{x \in U, \|x\|=1} \|\Omega x\| > t\sqrt{k} \right) \leq e^{-c_0 t^2 k}.$$

Proof. Let \mathcal{N} be a $(1/2)$ -net of the r -dimensional unit sphere of the image of U . Let \mathcal{M} be a $(1/2)$ -net of the k -dimensional unit sphere of the image of Ω . For any $u \in U$ where $\|u\|=1$, we can choose $x \in \mathcal{N}$ such that $\|x - u\|_2 < 1/2$. Then,

$$\|\Omega u\|_2 \leq \|\Omega x\|_2 + \|x - u\|_2 \max_{u_1 \in U, \|u_1\|=1} \|\Omega u_1\|.$$

Thus,

$$\max_{u_1 \in U, \|u_1\|=1} \|\Omega u_1\| \leq \|\Omega x\|_2 + \frac{1}{2} \max_{u_1 \in U, \|u_1\|=1} \|\Omega u_1\|.$$

This shows that $\|\Omega\| \leq 2 \sup_{x \in \mathcal{N}} \|\Omega x\|_2 = 2 \sup_{x \in \mathcal{N}} \sup_{v \in S^{k-1}} \langle \Omega x, v \rangle$. In a similar way, by approximating v with an element from \mathcal{M} we get

$$\sup_{x \in \mathcal{N}, v \in S^{k-1}} \langle \Omega x, v \rangle \leq \sup_{x \in \mathcal{N}, v \in \mathcal{M}} \langle \Omega x, v \rangle + \frac{1}{2} \sup_{x \in \mathcal{N}, v' \in S^{k-1}} \langle \Omega x, v' \rangle.$$

We obtain $\|\Omega\| \leq 4 \max_{x \in \mathcal{N}, y \in \mathcal{M}} |\langle \Omega x, y \rangle|$. By Lemma 2.1, we can choose these nets to be $|\mathcal{N}| \leq 6^r$ and $|\mathcal{M}| \leq 6^k$.

By Theorem 2.3, for every $x \in \mathcal{N}$ and $y \in \mathcal{M}$, the random variable $\langle \Omega x, y \rangle = \sum_{j=1}^k \sum_{k=1}^n a_{j,k} y_j x_k$ is sub-Gaussian, i.e. for $t > 0$

$$\mathbb{P}(|\langle \Omega x, y \rangle| > t\sqrt{k}) \leq C_2 e^{-c_1 t^2 k}.$$

By taking the union bound we get

$$\begin{aligned} \mathbb{P}(\|\Omega\|_2 > t\sqrt{k}) &\leq |\mathcal{N}| |\mathcal{M}| \mathbb{P}(|\langle \Omega x, y \rangle| > t\sqrt{k}/4, x \in \mathcal{N}, y \in \mathcal{M}) \\ &\leq 6^k \cdot 6^r \cdot C_2 e^{-c_1/16 t^2 k} \leq C_2 e^{-c_0 t^2 k}, \end{aligned}$$

provided that $t \geq C_0$ for an appropriately chosen constant $C_0 > 0$. This completes the proof. \square

By combining Theorem 3.4 with the ε -net argument and Lemmas 3.2 and 3.3 we obtain an estimate for the smallest value of $\|\Omega v\|$ for v in a subspace of dimension r .

Theorem 3.5 (Smallest value on a subspace). *There are constants M and D such that for any $n, r \in \mathbb{N}$, $p \in \mathbb{R}$, $0 < p < 1$, $n > r$, and for any r dimensional linear subspace $U \subset \mathbb{R}^n$, if $k > D \log \left(\frac{1}{p} \right) \left(r + \frac{1}{p^3} \right)$ then for $\Omega \in M_{k \times n}$ with centered sub-Gaussian random i.i.d entries as in Definition 2.4,*

$$\mathbb{P} \left(\min_{x \in U, \|x\|=1} \|\Omega x\|_2 \leq M\eta\sqrt{k} \right) \ll 1 \quad (3.1)$$

holds for $\eta < \mathcal{O}(\sqrt{p})$.

Proof. The proof is divided into three steps. In steps 1 and 2 Ω is bounded on incompressible and compressible vectors, respectively, and in step 3 these results are joined to complete the proof. We set M to be $M > \alpha C_0$ where C_0 comes from Theorem 3.4, and α from Lemma 3.2, such that $e^{-c_0 \frac{M}{\alpha} t^2 k}$ from Theorem 3.4 is sufficiently small.

Step 1: Let \mathcal{N} be a $\alpha\eta$ - net of the set of (ε_c, η) -incompressible vectors in the image of U . the number of vectors in \mathcal{N} is bounded by $\left(\frac{3}{\alpha\eta} \right)^r$. From Lemma 3.2 with $C = \frac{M}{\alpha}$ follows that for any vector $x \in \mathcal{N}$ and for $\varepsilon_c = \varepsilon_0 \eta \frac{1}{\sqrt{p}}$,

$$\mathbb{P}(\|\Omega x\|_2 < 2M\eta\sqrt{k}) \leq (C_1 \varepsilon_0 E(Z^3))^{k/4}.$$

Thus, by the union bound with failure probability of not more than

$$\left(\frac{3}{\alpha\eta}\right)^r \cdot (C_1\varepsilon_0 E(Z^3))^{k/4} \quad (3.2)$$

the following

$$\min_{x \in \mathcal{N}} \|\Omega x\|_2 \geq 2M\alpha\eta\sqrt{k} \quad (3.3)$$

holds. Since \mathcal{N} is an $\alpha\eta$ -net of the (ε_c, η) -incompressible vectors in the image of U , with the probability given in Eq. (3.2), then Eq.(3.3) holds. By Theorem 3.4 we have, for any incompressible vector y ,

$$\|\Omega y\| \geq \min_{x \in \mathcal{N}} \|\Omega x\|_2 - \alpha\eta\|\Omega\| \geq 2M\alpha\eta\sqrt{k} - M\alpha\eta\sqrt{k} = M\alpha\eta\sqrt{k}.$$

Step 2: Let \mathcal{M} be a η - net of the set of (ε_c, η) -compressible vectors in the image of U . The number of vectors in \mathcal{M} is bounded by Lemma 2.2 with $r^{\frac{1}{\varepsilon_c^2}} \eta^{-\frac{1}{\varepsilon_c^2}} \left(\frac{1}{\eta}\right)^r = r^{\frac{1}{\varepsilon_c^2}} \eta^{-\frac{1}{\varepsilon_c^2}}$. From Lemma 3.3 with $C = M$ it follows that for any vector $x \in \mathcal{M}$,

$$\mathbb{P}(\|\Omega x\|_2 < 2M\eta\sqrt{k}) \leq (1 - z'_4 p)^{k/4}.$$

Thus, by the union bound with failure probability of not more than

$$r^{\frac{1}{\varepsilon_c^2}} \eta^{-\frac{1}{\varepsilon_c^2}} \cdot (1 - z'_4 p)^{k/4}, \quad (3.4)$$

the following

$$\min_{x \in \mathcal{M}} \|\Omega x\|_2 \geq 2M\eta\sqrt{k} \quad (3.5)$$

holds. Since \mathcal{M} is an η -net of the (ε_c, η) -compressible vectors in the image of U , with the probability given in Eq.(3.4), then Eq. (3.5) holds. For any compressible vector y we have

$$\|\Omega y\| \geq \min_{x \in \mathcal{M}} \|\Omega x\|_2 - \eta\|\Omega\| \geq 2M\eta\sqrt{k} - M\eta\sqrt{k} = M\eta\sqrt{k}.$$

Thus, if Eqs. (3.4) and (3.2) are small enough, then, by Theorem 3.4

$$\min_{y \in U, \|y\|=1} \|\Omega y\| \geq M\alpha\varepsilon_c\sqrt{k}.$$

Step 3: The probabilities in Eqs. (3.2) and (3.4) are analyzed next. We have

$$\left(\frac{3}{\alpha\eta}\right)^r \cdot (C_1\varepsilon_0 E(Z^3))^{k/4} = e^{r \log(\frac{3}{\alpha\eta}) - k/4 \log(\frac{1}{C_1\varepsilon_0 z_3})}$$

$$r^{\frac{1}{\varepsilon_c^2}} \eta^{-\frac{1}{\varepsilon_c^2}} \cdot (1 - z'_4 p)^{k/4} \leq e^{\frac{1}{\varepsilon_0^2 \eta^2 p} \log(r) + \frac{1}{\varepsilon_0^2 \eta^2 p} \log(\frac{1}{\eta}) - c_1 p k}$$

for some c_1 that depends only on z_4 . Lemma 3.3 holds for $\eta = p^{1/2-\epsilon}$ and the probabilities in Eqs. (3.2) and (3.4) are less than

$$e^{r \log(\frac{3}{\alpha\varepsilon_0 \eta \sqrt{p}}) - k/4 \log(\frac{1}{M\varepsilon_0 z_3})} < e^{r \log(\frac{c_5}{p}) - c_6 k}$$

and

$$e^{\frac{1}{\varepsilon_0^2 \eta^2 p} \log(r) + \frac{1}{\varepsilon_0^2 \eta^2 p} \log(\frac{1}{\eta}) - c_1 p k} < e^{c_7 \frac{1}{p^2} \log(r) + c_8 \frac{1}{p^2} \log(\frac{1}{\eta}) - c_1 p k}$$

for constants c_i . Thus, for Eq. 3.1 to hold, k has to satisfy

$$c_9 r \log\left(\frac{c_5}{p}\right) \ll k \quad (3.6)$$

and

$$c_{10} \frac{1}{p^3} \log(r) + c_{11} \frac{1}{p^3} \log\left(\frac{1}{p}\right) \ll k. \quad (3.7)$$

Note that $\frac{1}{p^3} \log(r)$ is bounded by $\mathcal{O}(r \log(\frac{c_5}{p}))$ or $\mathcal{O}(\frac{1}{p^3} \log(\frac{1}{p}))$. Thus, there exists some constant D such that Eqs. 3.6 and 3.7 are equivalent to

$$D \log\left(\frac{1}{p}\right) \left(r + \frac{1}{p^3}\right) \ll k.$$

□

4 Approximated matrix decompositions

4.1 Randomized SVD using sparse projections

We present an algorithm that approximates the SVD decomposition of any matrix A . We first recall a known result (e.g. see [9]).

Theorem 4.1 (Theorem 11.2 in [9]). *Let A be an $m \times n$ matrix with singular values $\sigma_1, \dots, \sigma_n$ in descending order. For any integer $0 < r < m$, let Ω be a $n \times k$ random matrix. Denote $Y = A\Omega$ and $Y = QR$ where Q is a matrix with orthonormal columns and R is a full rank triangular matrix. If for any subspace $U \subset \mathbb{R}^n$ of dimension k , $\min_{x \in U} \|\Omega x\|_2$ and $\|\Omega\|_2$ are bounded from below and from above, respectively, with high probability, Then, with high probability,*

$$\|A - QQ^*A\|_2 \leq \mathcal{O}_\sigma(\sigma_{r+1}) \quad (4.1)$$

and

$$\|A - QQ^*A\|_F \leq \mathcal{O}_\sigma(\Delta_{r+1}). \quad (4.2)$$

Remark. Note that the notation $\mathcal{O}_\sigma(\sigma_{r+1})$ means that the error does not depend on the singular values except of having a linear dependency on σ_{r+1} . Dependency exists on n and k .

Remark 4.2. *Note that if $\Omega_1 \in M_{l \times n}$ and $\Omega_2 \in M_{k \times l}$ satisfy Eq. (4.1) for A of size $m \times n$ and $m \times l$ and $r \in \mathcal{N}$, respectively, then $\Omega = \Omega_2\Omega_1$ also satisfies Eq. (4.1) for $A \in M_{m \times n}$ and a rank r . This fact is important since it enables us to combine random matrices by utilizing for example a subsampled randomized Fourier transform (SRFT) [24] matrix or a Gaussian matrix with sub-Gaussian matrix. Similar statement is introduced in [3] as Fact 45.*

From Theorem 4.1 it follows that the randomized SVD Algorithm 5.1 in [9] is valid for sub-Gaussian matrices. This algorithm does not take advantage of the fact that Ω can be a sparse matrix. Thus, Algorithm 5.1 can be adapted similarly to the algorithm in Theorem 47 [3] and to the LU decomposition algorithm [1]. For SVD approximation to be of rank r , we use the following version of Weyl's inequality:

Theorem 4.3 (Weyl inequality for singular values). *Let $A, B \in M_{m \times n}$. If $\|A - B\|_2 \leq \varepsilon$, then for $1 \leq k \leq \min(m, n)$, $|\sigma_k(A) - \sigma_k(B)| \leq \varepsilon$ holds.*

Proof. We prove it by using the min-max principle that for any matrix $A \in M_{m \times n}$, if $\|A - B\|_2 \leq \varepsilon$ then $|\sigma_k(A) - \sigma_k(B)| \leq \varepsilon$.

The min-max principle states that

$$\sigma_k(A) = \max_{\substack{S \\ \dim S = n - k + 1}} \min_{\substack{x \in S, \\ \|x\| = 1}} \|Ax\|.$$

For any matrix S of dimension $n - k + 1$, we show that there exists a vector such that $\|Bx\| \leq \sigma_k(A) + \varepsilon$. For any such S , there is a vector $x \in S$ such that $\|Ax\| = \sigma_k(A)$. Note that

$$\varepsilon \geq \|A - B\| \geq \|(A - B)x\| \geq \|Ax\| - \|Bx\| = |\sigma_k(A) - \|Bx\||.$$

Thus, for any S of dimension $n - k + 1$, $\min_{x \in S, \|x\|=1} \|Bx\| \leq \sigma_k(A) + \varepsilon$. Therefore,

$$\sigma_k(B) \leq \sigma_k(A) + \varepsilon. \quad (4.3)$$

By repeating these considerations symmetrically for A with respect to B , we have that $\sigma_k(A) \leq \sigma_k(B) + \varepsilon$. Together with Eq. 4.3, we get $|\sigma_k(A) - \sigma_k(B)| \leq \varepsilon$. \square

Corollary 4.4. *If $\|A - U\Sigma V^*\|_2 \leq \mathcal{O}_\sigma(\sigma_{r+1}(A))$, then $\|A - U[\Sigma]_r V^*\|_2 \leq \mathcal{O}_\sigma(\sigma_{r+1}(A))$ where $[\Sigma]_r$ is the best rank r approximation of Σ .*

Proof.

$$\begin{aligned}
\|A - U[\Sigma]_r V^*\|_2 &= \|A - U\Sigma V^* + U\Sigma V^* - U[\Sigma]_r V^*\|_2 \\
&\leq \|A - U\Sigma V^*\|_2 + \|U\Sigma V^* - U[\Sigma]_r V^*\|_2 \\
&\leq O(\sigma_{r+1}(A)) + \sigma_{r+1}(B) \\
&\leq O(\sigma_{r+1}(A)) + \sigma_{r+1}(A) + O(\sigma_{r+1}(A)) \\
&= O(\sigma_{r+1}(A)).
\end{aligned}$$

□

Algorithm 4.1 describes a randomized SVD decomposition for getting a rank r approximation. This approximation generates the error $\mathcal{O}_\sigma(\sigma_{r+1}(A))$. Theorem 4.5 proves that the algorithm is correct for any matrix distribution that holds the conditions of Theorem 4.1. Its complexity is evaluated in section 4.1.1. Numerical results are given in section 5.

Algorithm 4.1: Sub-Gaussian-based Randomized SVD Decomposition

Input: A matrix of size $m \times n$ to decompose, r desired rank, k_1, k_2, l number of columns to use.

Output: Matrices U, Σ, V such that $\|A - U\Sigma V^*\|_2 \leq \mathcal{O}_\sigma(\sigma_{r+1}(A))$ where U and V are matrices with orthonormal columns.

- 1: Create a random sub-Gaussian matrix Ω_1 of size $k_1 \times n$.
 - 2: Create a random Gaussian matrix Ω'_1 of size $l \times k_1$.
 - 3: Compute $B = A\Omega_1^* \Omega'^*_1$ ($B \in M_{m \times l}$).
 - 4: Compute the QR decomposition: $B = QR$, $Q \in M_{m \times k_1}$ with orthonormal columns, $R \in M_{k_1 \times k_1}$ is a full rank upper triangular matrix.
 - 5: Create a random sub-Gaussian matrix Ω_2 of size $k_2 \times m$.
 - 6: Compute $\Omega_2 Q$, $\Omega_2 A$ and $(\Omega_2 Q)^\dagger$.
 - 7: Compute the SVD of $(\Omega_2 Q)^\dagger \Omega_2 A = \tilde{U}_1 \Sigma_1 V_1^*$.
 - 8: $\tilde{U} \leftarrow \tilde{U}_1(:, 1:r)$.
 - 9: $\Sigma \leftarrow \Sigma_1(1:r, 1:r)$.
 - 10: $V \leftarrow V_1(:, 1:r)$.
 - 11: $U \leftarrow Q\tilde{U}$.
-

Theorem 4.5. Assume that A is a matrix of size $m \times n$ where $m < n$ and $r < m$. Then for $k_1, k_2 = \mathcal{O}\left(\log\left(\frac{1}{p}\right)\left(r + \frac{1}{p^3}\right)\right)$ and $l = \mathcal{O}(r)$, Algorithm 4.1 outputs U, Σ and V such that $\|A - U\Sigma V^*\|_2 \leq \mathcal{O}_\sigma(\sigma_{r+1}(A))$.

Proof. For a matrix $A \in M_{m \times n}$, let $\Omega_1 \in M_{k_1 \times n}$ be a sub-Gaussian matrix and let $\Omega'_1 \in M_{l \times k_1}$ be a random Gaussian matrix. Denote the QR-decomposition of $A\Omega_1^* \Omega'^*_1 \in M_{m \times l}$ by $QR = A\Omega_1^* \Omega'^*_1$. From Theorem 4.1, Remark 4.2 and Theorem 10.8 of [9], it follows that for $l = \mathcal{O}(r)$

$$\|QQ^*A - A\|_2 \leq \mathcal{O}_\sigma(\sigma_{r+1}) \quad (4.4)$$

holds. From Theorem 3.5 it follows that for a sub-Gaussian matrix $\Omega_2 \in M_{k_2 \times m}$, where $k_2 = \mathcal{O}(k_1)$, the matrix $\Omega_2 Q$ is invertible from the left, namely $(\Omega_2 Q)^\dagger \Omega_2 Q = I_{k_1 \times k_1}$. Thus, $\|QQ^*A - A\|_2 = \|Q(\Omega_2 Q)^\dagger (\Omega_2 Q) Q^* A - A\|_2$. From the construction $\|U_1 \Sigma_1 V_1^* - A\|_2 = \|Q\tilde{U} \Sigma V^* - A\|_2 = \|Q(\Omega_2 Q)^\dagger \Omega_2 A - A\|_2$. $\|Q(\Omega_2 Q)^\dagger \Omega_2 A - A\|_2$ is bounded in the following way:

$$\begin{aligned}
\|Q(\Omega_2 Q)^\dagger \Omega_2 A - A\|_2 &\leq \|Q(\Omega_2 Q)^\dagger \Omega_2 A - Q(\Omega_2 Q)^\dagger (\Omega_2 Q) Q^* A \\
&\quad + Q(\Omega_2 Q)^\dagger (\Omega_2 Q) Q^* A - A\|_2 \\
&= \|Q(\Omega_2 Q)^\dagger \Omega_2 (A - QQ^*A) + QQ^*A - A\|_2 \\
\text{by the triangle inequality} &\leq \|Q(\Omega_2 Q)^\dagger \Omega_2 (A - QQ^*A)\|_2 + \|QQ^*A - A\|_2 \\
\text{since } \|AB\|_2 &\leq \|A\|_2 \|B\|_2 &\leq \|Q(\Omega_2 Q)^\dagger \Omega_2\|_2 \|A - QQ^*A\|_2 + \|QQ^*A - A\|_2 \\
&= ((\|Q(\Omega_2 Q)^\dagger \Omega_2\|_2 + 1) \|A - QQ^*A\|_2 \\
\text{since } \|AB\|_2 &\leq \|A\|_2 \|B\|_2 &\leq ((\|Q(\Omega_2 Q)^\dagger\|_2 \|\Omega_2\|_2 + 1) \|A - QQ^*A\|_2. \quad (4.5)
\end{aligned}$$

From Theorem 3.5 we have $\|(\Omega_2 Q)^\dagger\|_2 \leq 1/(c_1 \sqrt{k_2})$. From Theorem 2.4 we have $\|\Omega_2\|_2 \leq C_0 \sqrt{n}$. Thus, from Eq. (4.5) it follows that

$$\|Q(\Omega_2 Q)^\dagger \Omega_2 A - A\|_2 \leq \left(\frac{C_0}{c_1} \sqrt{\frac{n}{k_2}} + 1 \right) \|A - Q Q^* A\|_2.$$

Together with Eq. (4.4) we get that $\|U_1 \Sigma_1 V_1^* - A\|_2 \leq \mathcal{O}_\sigma(\sigma_{r+1})$.

From the result of Corollary 4.4 we get $\|U \Sigma V^* - A\|_2 \leq \mathcal{O}_\sigma(\sigma_{r+1})$. \square

Remark. A bound for the Frobenius norm $\|A - U \Sigma V^*\|_F \leq \mathcal{O}_\sigma(\Delta_{r+1}(A))$ is reached similarly by using Eq. (4.2).

4.1.1 Computational Complexity of Algorithm 4.1

For computational complexity estimation and implementation, the internal random matrix distribution of the algorithm is selected as a subclass of sparse sub-Gaussian matrices. We chose sparse-Gaussian matrices. Sparse-Gaussian matrices are sparse matrices, where each entry is i.i.d with probability $1 - p$ to be zero and standard Gaussian otherwise. The complexity of each step in Algorithm 4.1 is shown in Table 4.1

Table 4.1: Complexity of Algorithm 4.1

Step in Algorithm 4.1	A sparse	A dense
Creation of sparse matrix Ω_1 of size $k_1 \times n$	$\mathcal{O}(n)$	$\mathcal{O}(n)$
Computation of $B = A \Omega_1^*$	$\mathcal{O}(\text{nnz}(A) p k_1 + m k_1 l)$	$\mathcal{O}(m n p k_1 + m k_1 l)$
Computation of its QR- decomposition, $B = QR$	$\mathcal{O}(m k_1^2)$	$\mathcal{O}(m k_1^2)$
Creation of sparse matrix Ω_2 of size $k_2 \times m$	$\mathcal{O}(m)$	$\mathcal{O}(m)$
Computation of $\Omega_2 Q$, $\Omega_2 A$	$\mathcal{O}(m k_2 + \text{nnz}(A) p k_2)$	$\mathcal{O}(m k_2 + m n p k_2)$
Computation of $(\Omega_2 Q)^\dagger \Omega_2 A$	$\mathcal{O}(k_1 k_2^2 + n k_2)$	$\mathcal{O}(k_1 k_2^2 + n k_2)$
Computation of the SVD of $(\Omega_2 Q)^\dagger \Omega_2 A$	$\mathcal{O}(n k_2^2)$	$\mathcal{O}(n k_2^2)$

The total complexity is $\mathcal{O}(\text{nnz}(A) p k + (m + n) k^2)$. For example, for sub-Gaussian random matrices with $p = \mathcal{O}(\frac{1}{\sqrt[3]{r}})$ and $k = \mathcal{O}(r \log r)$, the complexity is $\mathcal{O}(\text{nnz}(A) \sqrt[3]{r^2} \log r + (m + n)(r \log r)^2)$.

For the OSE defined in [15], the asymptotic complexity is the same as in [15]. We show in Section 5 that although the asymptotic complexity is the same, Algorithm 4.1 is faster in practice.

4.2 Sub-Gaussian based Randomized LU decomposition

Theorem 4.1 is equivalent to Theorem 3.1 in [1] where L_2 norm is used instead of Frobenius norm. A sub-Gaussian distribution can be used instead of the sparse embedding matrix distribution. Since the correctness proof of the algorithm in [1] is based on Theorem 3.1, it is also applicable for sub-Gaussian matrices.

Theorem 4.6. *Assume that sub-Gaussian random matrices are used instead of sparse embedding matrices in the approximated rank r LU decomposition in [1]. Then, for any $r \in \mathcal{N}$, and for any matrix $A \in M_{m \times n}$, the approximated rank r LU decomposition results in matrices L and U and permutations P and Q such that $\|PAQ - LU\|_2 \leq \mathcal{O}_\sigma(\sigma_{r+1}(A))$.*

The complexity of the algorithm, as shown in [1], is $\mathcal{O}(\text{nnz}(A) p k + (m + n) k^2)$.

5 Numerical Results

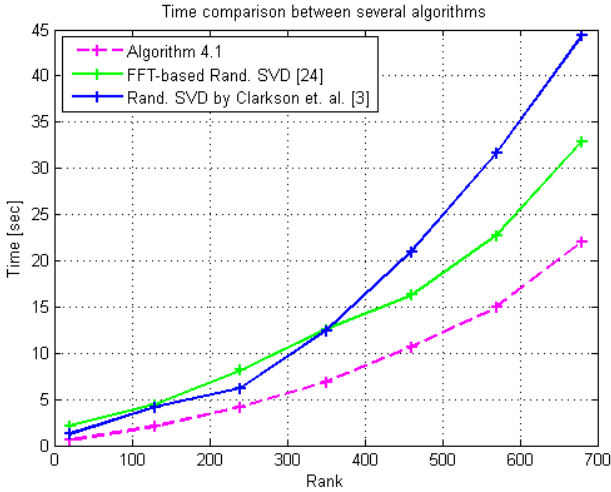
The results in this paper are valid to all types of i.i.d sub-Gaussian matrices and OSE distributions. In the current implementation, we used sparse-Gaussian matrices, where each entry in the matrix is i.i.d with probability p to be standard Gaussian and zero otherwise. Note that this distribution is like the distribution in Definition 2.4 up to a multiplicative constant that does not affect Algorithm 4.1.

We noticed that in practice, for the decomposition of specific matrices, the use of internal random matrix distribution, which has more than one non-zero entry in a each row (as presented for example here and in [15]), results in a much better approximation error than distributions with one non-zero in each row as in [3]. This is the reason we use $p = 3/n$ in the sparse-Gaussian matrices implementation.

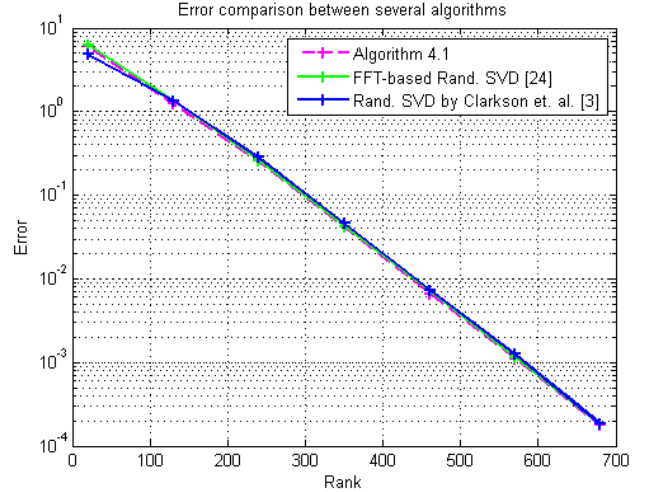
We describe the results from three different experiments. All the experiments were implemented on Intel Xeon CPU X5560 2.8GHz. All the experiments compare between the running time and the generated error from the following three algorithms in different scenarios: 1. The FFT-based algorithm given in [24]. 2. The Algorithm from [3]. 3. Algorithm 4.1. Although the proven error bounds for Algorithm 4.1 are less tight than the bounds for the other algorithms, we see that in practice Algorithm 4.1 reaches the same error. In all the experiments, the parameters for the different algorithms are chosen such that the reconstruction error rates are similar and aligned to the error from [3] and [24]. The slowest algorithm has an error that is not smaller than the fastest algorithm.

The experiments that took place are:

1. Rank r approximation is computed for a randomly generated full matrix $A \in M_{3000 \times 3000}$ with singular values that decay exponentially fast from 1 to e^{-50} . Figure 5.1 displays the comparison between the running time and the error from rank r approximation from the three algorithms mentioned above. The x-axis denotes the rank and the y-axis denotes the running time. The results show that for a small rank range [3] is faster than the FFT-based algorithm [24]. For a larger rank range, the FFT-based algorithm is faster. For all ranks, Algorithm 4.1 is the fastest.



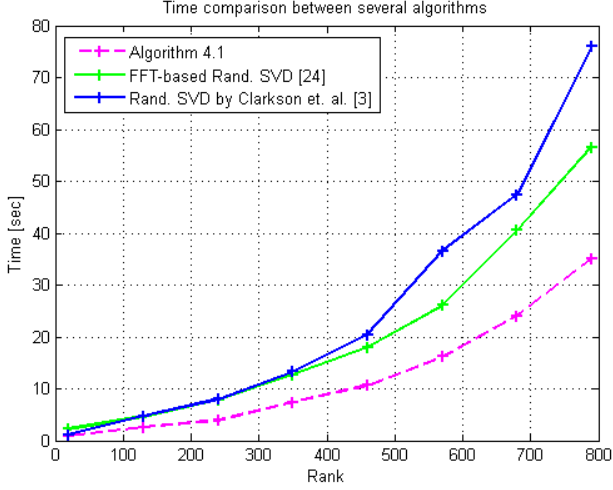
(a) Time



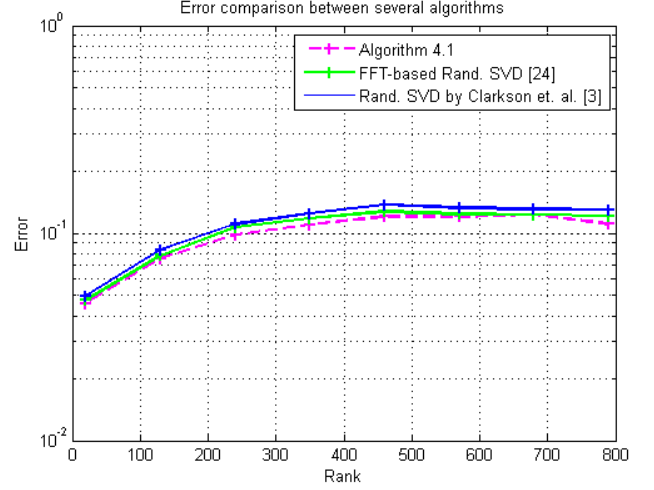
(b) Error

Figure 5.1: Results from the approximation of a matrix of size 3000×3000 with exponentially decaying singular values. The x-axes in both (a) and (b) denote the rank of the approximation. The y-axis in (a) denotes the run time. The y-axis in (b) denotes the error from the rank approximation.

2. Rank r approximation is computed for a randomly generated full matrix $A \in M_{3000 \times 3000}$ where the first r singular values are 1 and the other singular values decay exponentially fast from e^{-5} to e^{-50} . Figure 5.2a displays the comparison between the running time for rank r approximation for the three algorithms mentioned above. x-axis denotes the rank and the y-axis denotes the running time. As in experiment 1, for a small rank range, [3] is faster than the FFT-based algorithm [24]. For a larger rank range, the FFT-based algorithm is faster than [3]. For all ranks, Algorithm 4.1 is the fastest.



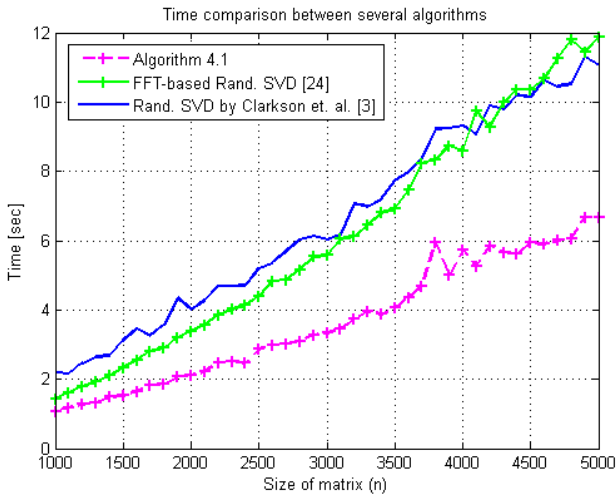
(a) Time



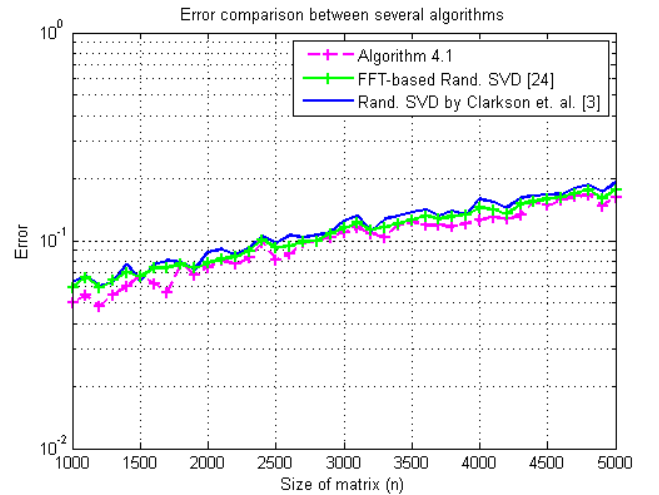
(b) Error

Figure 5.2: Results from the approximation of a matrix of size 3000×3000 with different numerical ranks. The x-axes in both (a) and (b) denote the numerical rank. The y-axis in (a) denotes the run time. The y-axis in (b) denotes the rank approximation error.

- Rank 300 approximation of a randomly generated full matrix $A \in M_{n \times n}$ is computed when the first 300 singular values are 1 and the other singular values decay exponentially fast from e^{-5} to e^{-50} . Figure 5.3 displays the comparison between the run time for rank 300 approximation from the three algorithms mentioned above. x-axis denotes the rank and y-axis denotes the running time. It is noticeable in this experiment that the sparse SVD from [3] is faster than the FFT-based algorithm [24] when n increases. For rank 300 and for $n \approx 4500$ the algorithm from [3] is faster than the FFT-based algorithm. For ranks larger than 300, a large n is required for the algorithm from [3] to be faster than the FFT-based algorithm. The Sparse SVD Algorithm 4.1 presented in this paper is faster for all n .



(a) Time



(b) Error

Figure 5.3: Results from the approximation a matrix of size $n \times n$, $n = 1000, \dots, 5000$ with numerical rank 300. The x-axis in both (a) and (b) denotes n . The y-axis in (a) denotes the run time. The y-axis in (b) denotes the approximation error.

In Algorithm 4.1, it is only necessary to apply the matrix A once from the left and once from the right, then A does not have to be stored in memory. Table 5.1 shows the running time for large

matrices that cannot be stored in a computer memory. The matrices we chose have a similar form to the choice in [8]. We chose $A = F\Sigma F$ where F is the DFT matrix and Σ is a diagonal matrix with singular values σ_i that decay linearly until $i = 200$ and exponentially from there on. We set $\sum_{i=201}^n \sigma_i$ to be constant in this experiment. Algorithm 4.1 is applied to rank 200 with $k_1 = 500$ and $k_2 = 700$.

Size (n)	Relative Error from Algorithm 4.1	Time for Alg. 4.1(sec)	Time for full SVD
1,024	1.5465	1.0011	1.5232
2,048	1.5645	1.6236	11.3702
4,096	1.5422	2.7653	94.6345
8,192	1.5571	5.2999	578.2982
16,384	1.4846	12.1065	4324.683
32,768	1.5686	26.4022	
65,536	1.5074	50.6191	
131,072	1.4838	109.8185	
262,144	1.5357	205.0068	
524,288	1.4854	418.4137	
1,048,576	1.5240	847.8211	

Table 5.1: Comparing running time of Algorithm 4.1 to the standard SVD for large matrices. The Relative Error is the ratio between the error from the rank r decomposition and from the $r + 1$ singular value.

Conclusion

We showed that matrices with i.i.d sub-Gaussian entries conserve subspaces and showed the connection between the distribution of the entries and the required size of the matrix. A new algorithm is presented, which yields with high probability, a rank r SVD approximation for an $m \times n$ matrix that achieves an asymptotic complexity of $\mathcal{O}(\text{nnz}(A)pk + (m + n)k^2)$. Additionally, we showed that the approximated LU algorithm in [1], which uses sub-Gaussian random matrices, has a computational complexity of $\mathcal{O}(\text{nnz}(A)pk + (m + n)k^2)$. We showed in the experiments that although the derived error bounds are not as tight as the bounds from the algorithms in [3, 9], in practice, the algorithm in this paper reaches the same error in less time.

Future work includes non-asymptotic estimation of the algorithm parameters including error estimation improvement to get tighter bounds.

Acknowledgment

This research was partially supported by the Israeli Ministry of Science & Technology (Grants No. 3-9096, 3-10898), US-Israel Binational Science Foundation (BSF 2012282), Blavatnik Computer Science Research Fund, Blavatnik ICRC Funds and by a Fellowship from Jyväskylä University. We thank Prof. Jelani Nelson and Dr. Haim Avron for their continues support and constructive remarks.

References

- [1] Y. AIZENBUD, G. SHABAT, AND A. AVERBUCH, *Randomized LU decomposition using sparse projections*, Preprint, (2015).
- [2] A. C. BERRY, *The accuracy of the gaussian approximation to the sum of independent variates*, Transactions of the american mathematical society, 49 (1941), pp. 122–136.
- [3] K. L. CLARKSON AND D. P. WOODRUFF, *Low rank approximation and regression in input sparsity time*, in Proceedings of the 45th annual ACM symposium on Symposium on theory of computing, ACM, 2013, pp. 81–90.

- [4] M. B. COHEN, *Nearly tight oblivious subspace embeddings by trace inequalities*, in Proceedings of the Twenty-Seventh Annual ACM-SIAM Symposium on Discrete Algorithms, 2016, pp. 278–287.
- [5] A. DASGUPTA, R. KUMAR, AND T. SARLÓS, *A sparse johnson: Lindenstrauss transform*, in Proceedings of the forty-second ACM symposium on Theory of computing, ACM, 2010, pp. 341–350.
- [6] S. DIRKSEN, *Dimensionality reduction with subgaussian matrices: a unified theory*, arXiv preprint arXiv:1402.3973, (2014).
- [7] C.-G. ESSEEN, *On the Liapounoff limit of error in the theory of probability*, Almqvist & Wiksell, 1942.
- [8] N. HALKO, P.-G. MARTINSSON, Y. SHKOLNISKY, AND M. TYGERT, *An algorithm for the principal component analysis of large data sets*, SIAM Journal on Scientific computing, 33 (2011), pp. 2580–2594.
- [9] N. HALKO, P.-G. MARTINSSON, AND J. A. TROPP, *Finding structure with randomness: Probabilistic algorithms for constructing approximate matrix decompositions*, SIAM review, 53 (2011), pp. 217–288.
- [10] W. Hoeffding, *Probability inequalities for sums of bounded random variables*, Journal of the American statistical association, 58 (1963), pp. 13–30.
- [11] W. B. JOHNSON AND J. LINDENSTRAUSS, *Extensions of lipschitz mappings into a hilbert space*, Contemporary mathematics, 26 (1984), p. 1.
- [12] D. M. KANE AND J. NELSON, *Sparser johnson-lindenstrauss transforms*, Journal of the ACM (JACM), 61 (2014), p. 4.
- [13] A. E. LITVAK, A. PAJOR, M. RUDELSON, AND N. TOMCZAK-JAEGERMANN, *Smallest singular value of random matrices and geometry of random polytopes*, Advances in Mathematics, 195 (2005), pp. 491–523.
- [14] P.-G. MARTINSSON, V. ROKHLIN, AND M. TYGERT, *A randomized algorithm for the decomposition of matrices*, Applied and Computational Harmonic Analysis, 30 (2011), pp. 47–68.
- [15] J. NELSON AND H. L. NGUYÊN, *Osnap: Faster numerical linear algebra algorithms via sparser subspace embeddings*, in Foundations of Computer Science (FOCS), 2013 IEEE 54th Annual Symposium on, IEEE, 2013, pp. 117–126.
- [16] ———, *Sparsity lower bounds for dimensionality reducing maps*, in Proceedings of the forty-fifth annual ACM symposium on Theory of computing, ACM, 2013, pp. 101–110.
- [17] M. RUDELSON, *Recent developments in non-asymptotic theory of random matrices*, Modern Aspects of Random Matrix Theory, 72 (2014), p. 83.
- [18] M. RUDELSON AND R. VERSHYNIN, *The littlewood–offord problem and invertibility of random matrices*, Advances in Mathematics, 218 (2008), pp. 600–633.
- [19] ———, *Smallest singular value of a random rectangular matrix*, Communications on Pure and Applied Mathematics, 62 (2009), pp. 1707–1739.
- [20] G. SHABAT, Y. SHMUELI, Y. AIZENBUD, AND A. AVERBUCH, *Randomized LU decomposition*, arXiv preprint arXiv:1310.7202, (2013).
- [21] T. TAO, *Topics in random matrix theory*, vol. 132, American Mathematical Soc., 2012.
- [22] J. A. TROPP, *Improved analysis of the subsampled randomized hadamard transform*, Advances in Adaptive Data Analysis, 3 (2011), pp. 115–126.
- [23] R. VERSHYNIN, *Introduction to the non-asymptotic analysis of random matrices*, arXiv preprint arXiv:1011.3027, (2010).

- [24] F. WOOLFE, E. LIBERTY, V. ROKHLIN, AND M. TYGERT, *A fast randomized algorithm for the approximation of matrices*, Applied and Computational Harmonic Analysis, 25 (2008), pp. 335–366.